CrossMark

# A New Family of Generalized Distributions Based on Alpha Power Transformation with Application to Cancer Data

M. Nassar[1] · A. Alzaatreh[2] · O. Abo-Kasem[1] ·
M. Mead[1] · M. Mansoor[3]

**Abstract** In this paper, we propose a new method for generating distributions based on the idea of alpha power transformation introduced by Mahdavi and Kundu (Commun Stat Theory Methods 46(13):6543–6557, 2017). The new method can be applied to any distribution by inverting its quantile function as a function of alpha power transformation. We apply the proposed method to the Weibull distribution to obtain a three-parameter alpha power within Weibull quantile function. The new distribution possesses a very flexible density and hazard rate function shapes which are very useful in cancer research. The hazard rate function can be increasing, decreasing, bathtub or upside down bathtub shapes. We derive some general properties of the proposed distribution including moments, moment generating function, quantile and Shannon entropy. The maximum likelihood estimation method is used to estimate the parameters. We illustrate the applicability of the proposed distribution to complete and censored cancer data sets.

**Keywords** Alpha power transformation · Moments · Weibull distribution · Maximum likelihood estimation · Shannon entropy

✉ A. Alzaatreh
aalzaatreh@aus.edu

[1] Department of Statistics, Faculty of Commerce, Zagazig University, Zagazig, Egypt

[2] Department of Mathematics and Statistics, American University of Sharjah, Sharjah, UAE

[3] Department of Statistics, The Islamia University of Bahawalpur, Bahawalpur, Pakistan

## 1 Introduction

The idea of developing new distributions remains an important topic in the recent literatures. It provides more flexible distributions that can model complex data structure. Lee et al. [12] in their review paper provided an overview of most methods used to generate family of continuous distributions. They pointed out that prior to 1980, methods of generating distributions can be categorized into three categories; method of differential equation, method of transformation and method of quantile function. For more details about these methods, one is referred to Pearson [20], Johnson [10] and Tukey [23]. Since 1980, several methods of generating distributions proposed in the literature. Lee et al. [12] categorized these methods as "method of combination". These methods focused mainly on adding parameters to an existing distribution or combining existing distributions. For more details about the recent developments in generalizing distributions, we refer the reader to Johnson et al. [11], Eugene et al. [7], Jones [9], Alzaatreh et al. [1–3] and Tahir et al. [22].

Recently, Mahdavi and Kundu [16] proposed the so called alpha power transformation (APT) family. The parameter $\alpha$ is introduced to incorporate skewness to the base distribution. The APT family is defined as follows: Let $F(x)$ be the cumulative density function (CDF) of any continuous random variable $X$, then the CDF of the APT family is given by

$$F_{\text{APT}}(x; \alpha) = \begin{cases} \frac{\alpha^{F(x)} - 1}{\alpha - 1} & \text{if } \alpha > 0, \alpha \neq 1 \\ F(x) & \text{if } \alpha = 1. \end{cases} \tag{1}$$

The corresponding probability density function (PDF) is

$$f_{\text{APT}}(x; \alpha) = \begin{cases} \frac{\log \alpha}{\alpha - 1} f(x) \alpha^{F(x)} & \text{if } \alpha > 0, \alpha \neq 1 \\ f(x) & \text{if } \alpha = 1. \end{cases}$$

Mahdavi and Kundu [16] applied the proposed method to the exponential distribution and proposed the alpha power exponential distribution. They studied various properties of the proposed distribution such as explicit expressions for the moments, quantiles and moment generating function.

This paper is organized in the following way: In Sect. 2, we propose a new method for generating continuous distributions based on the family of distributions in (1). The proposed family of distributions has a connection with weighted distributions. In Sect. 3, a member of the proposed family namely, Alpha Power within Weibull Quantile Distribution (APWQ), is proposed. General properties of the APWQ are studied in Sect. 4 including, quantile, moments, moment generating function, Shannon entropy, mean residual life and mean waiting time functions. The maximum likelihood estimation and Applications to complete and censored cancer data sets are studied in Sect. 5. Section 6 offers some concluding remarks.

## 2 General Properties of the New Method

Let $g(x)$ and $G(x)$ be, respectively, the PDF and CDF of any random variable $X$. Then the CDF, $F(x)$, of the new proposed method for generating distributions can be obtained by inverting the following equation

$$\frac{\alpha^{F(x)} - 1}{\alpha - 1} = G(x), \alpha \neq 1. \tag{2}$$

Therefore,

$$F(x) = \frac{\log(1 + (\alpha - 1)G(x))}{\log(\alpha)}, x \in \mathbb{R}, \quad \alpha > 0, \alpha \neq 1. \tag{3}$$

The corresponding PDF is

$$f(x) = \frac{(\alpha - 1)g(x)}{\log(\alpha)(1 + (\alpha - 1)G(x))} \tag{4}$$

Note that when $\alpha \to 1$, $f(x)$ reduces to $g(x)$. Equation (4) can be written in the following form

$$f(x) = \frac{g(x)\omega(x)}{c}. \tag{5}$$

From (5), it is clear that $f(x)$ is a weighted version of $g(x)$, where the weight function is

$$\omega(x) = (1 + (\alpha - 1)G(x))^{-1}$$

and the normalizing constant $c = \log(\alpha)/(\alpha - 1)$. A useful expansion for the CDF and PDF in (3) and (4) for $0 \leq \alpha \leq 2, \alpha \neq 1$ are given by

$$F(x) = \frac{1}{\log(\alpha)} \sum_{k=1}^{\infty} \frac{(-1)^k(\alpha - 1)^k}{k} G(x)^k.$$

and

$$f(x) = \frac{g(x)}{\log(\alpha)} \sum_{k=0}^{\infty} (-1)^k(\alpha - 1)^{k+1} G(x)^k \tag{6}$$

From (3) and (4), the hazard rate function, $h(x)$, is given by

$$h(x) = \frac{(\alpha - 1)g(x)}{(1 + (\alpha - 1)G(x)) \left[\log(\alpha) - \log(1 + (\alpha - 1)G(x))\right]}.$$

*Remark 1* If $X$ follows the distribution in (3), then the quantile function is given by

$$x_q = G^{-1}\left(\frac{\alpha^q - 1}{\alpha - 1}\right), 0 \leq q \leq 1.$$

Note that Remark 1 can be used to simulate random sample from $F(x)$ distribution by first simulating random sample $U_i \sim \text{Uniform}(0, 1), i = 1, \ldots, n$. Then the random sample $X_i = G^{-1}\left(\frac{\alpha^{U_i}-1}{\alpha-1}\right), i = 1, \ldots, n$ follow $F(x)$ distribution.

**Theorem 1** *If a random variable X follows the family of distributions in* (4)*, then the Shannon entropy defined as $\eta_X = E\left[-\log f(x)\right]$ is given by*

$$\eta_X = \log\left(\frac{\sqrt{\alpha}\log(\alpha)}{\alpha-1}\right) - E\{\log g(X)\}. \tag{7}$$

## 3 Alpha Power within Weibull Quantile Distribution

The Weibull distribution is a popular life time distribution in reliability theory. Numerous articles have been written demonstrating the applications of the Weibull distribution in biology, medicine, engineering and meteorology. In the last few years, several researchers have developed various extensions and generalizations of the Weibull distribution to model various types of data. Among these, Mudholkar et al. [18,19] introduced and studied the exponentiated Weibull distribution to analyze bathtub failure data by adding an extra shape parameter to the Weibull distribution. Xie and Lai [24] introduced the additive Weibull distribution, Jalmar et al. [8] introduced the generalized modified Weibull distribution and Cordeiro et al. [5] studied the exponential-Weibull distribution. Next, Eq. (3) is used to introduce the APWQ distribution.

Let $X$ be a random variable follows the Weibull distribution with CDF $G(x) = 1 - e^{-\lambda x^\beta}, x > 0$. From (3), the CDF of the APWQ distribution is defined as

$$F(x) = \frac{\log(1 + (\alpha-1)(1 - e^{-\lambda x^\beta}))}{\log(\alpha)}, x > 0. \tag{8}$$

The corresponding PDF is

$$f(x) = \frac{(\alpha-1)\lambda\beta x^{\beta-1}e^{-\lambda x^\beta}}{\log(\alpha)(1 + (\alpha-1)(1 - e^{-\lambda x^\beta}))}, \tag{9}$$

where $\alpha, \beta > 0$ are shape parameters and $\lambda > 0$ is a scale parameter. Table 1 lists various special models of the APWQ distribution.

*Remark 2* Using the result in (6), the PDF in (9) for $0 \leq \alpha \leq 2, \alpha \neq 1$, can be expressed in a generalized mixture form of the Weibull distributions as

$$f(x) = \frac{1}{\log(\alpha)}\sum_{k=0}^{\infty}\sum_{j=0}^{k}\binom{k}{j}(-1)^{k+j}(\alpha-1)^{k+1}g_{WD}(x; \lambda(j+1), \beta) \tag{10}$$

where $g_{WD}(x; \lambda(j+1), \beta)$ is the PDF of the Weibull distribution with scale parameter $\lambda(j+1)$ and shape parameter $\beta$.

**Table 1** Sub-models of the APWQ distribution

| $\alpha$ | $\lambda$ | $\beta$ | Reduced model |
| --- | --- | --- | --- |
| 1 | – | – | Two parameter Weibull distribution |
| – | – | 2 | Alpha power within Rayleigh quantile distribution |
| 1 | – | 2 | Rayleigh distribution |
| – | – | 1 | Alpha power within exponential quantile distribution |
| 1 | – | 1 | Exponential distribution |

On the other hand, if $\alpha > 2$, and using the expansion

$$\left[1 + (\alpha - 1)(1 - e^{-\lambda x^\beta})\right]^{-1} = \alpha^{-1} \sum_{k=0}^{\infty} (1 - 1/\alpha)^k e^{-\lambda k x^\beta}$$

the PDF in (9) can be written as

$$f(x) = \frac{1}{\log(\alpha)} \sum_{k=0}^{\infty} (1 - 1/\alpha)^{k+1} g_{WD}(x; \lambda(k+1), \beta) \tag{11}$$

The survival function and hazard rate function for APWQ are, respectively, given by

$$S(x) = \frac{\log(\alpha) - \log(1 + (\alpha - 1)(1 - e^{-\lambda x^\beta}))}{\log(\alpha)},$$

and

$$h(x) = \frac{(\alpha - 1)\lambda \beta x^{\beta-1} e^{-\lambda x^\beta}}{(1 + (\alpha - 1)(1 - e^{-\lambda x^\beta})) \left[\log(\alpha) - \log(1 + (\alpha - 1)(1 - e^{-\lambda x^\beta}))\right]}, x > 0.$$

Figures 1 and 2 display some plots of the APWQ density and hazard rate functions respectively for various parameter values of $\alpha$ and $\beta$ where the scale parameter $\lambda = 1$. These plots show that the APWQ is flexible in terms of shapes. The APWQ distribution can be left-skewed or right-skewed. Also, the hazard rate function can be very flexible. It can be increasing (IFR), decreasing (DFR), bathtub (BT), upside down bathtub (UBT) or bimodal failure rate shapes.

## 4 Properties of the APWQ Distribution

In this section, we provide some general properties of the APWQ distribution including quantile function, mode, moments, entropy, order statistics and mean residual life and mean waiting time.
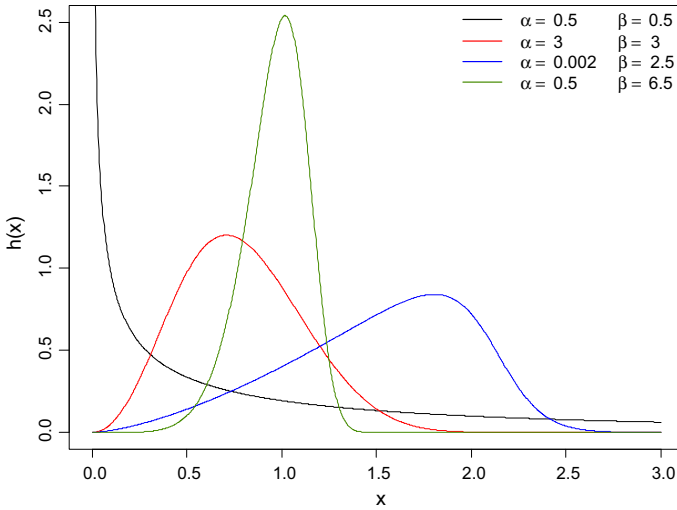
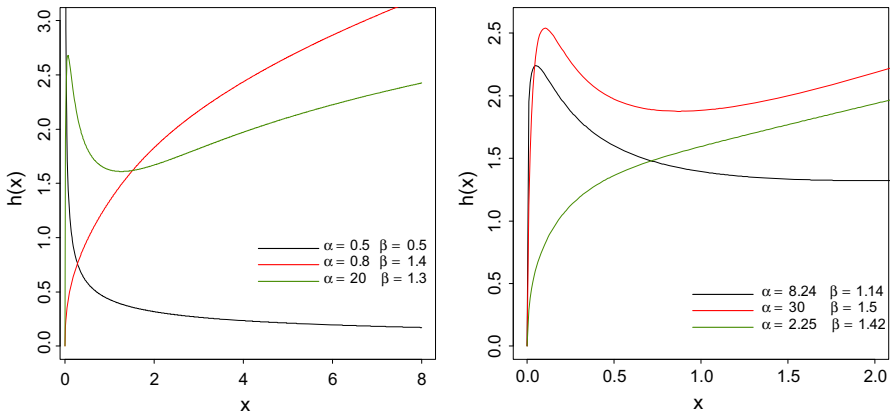**Fig. 1** Density plots of APWQ density for various values of $\alpha$ and $\beta$



**Fig. 2** Hazard rate plots of APWQ distribution for various values of $\alpha$ and $\beta$

*Remark 3* The $q$-th quantile function of APWQ distribution is given by

$$x_q = \left( \frac{-1}{\lambda} \log \left[ \frac{\alpha - \alpha^q}{\alpha - 1} \right] \right)^{1/\beta}, 0 \leq q \leq 1.$$

**Theorem 2** *The APWQ is unimodal. When $\beta \leq 1$, the mode is at $x = 0$. And when $\beta > 1$, the mode is at $x = x_0$ where $k(x_0) = 0$ and*

$$k(x) = 1 - \beta + \alpha \left\{ \beta - 1 + e^{x^\beta \lambda}(1 - \beta + x^\beta \beta \lambda) \right\}.$$

*Proof* Since $\lambda$ is a scale parameter, without loss of generality assume $\lambda = 1$. From (9), $f'(x) = 0 \Leftrightarrow x^{\beta-2}[(\beta - 1 - \lambda\beta x^{\beta})(\alpha + (1-\alpha)e^{-\lambda x^{\beta}}) + \lambda\beta(1-\alpha)x^{\beta}e^{-\lambda x^{\beta}}] = 0$. Therefore the critical values of $f(x)$ are $x = 0$ or the solution of the equation $(\beta - 1 - \lambda\beta x^{\beta})(\alpha + (1-\alpha)e^{-\lambda x^{\beta}}) + \lambda\beta(1-\alpha)x^{\beta}e^{-\lambda x^{\beta}} = 0$. This implies that $(\beta - 1 - \lambda\beta x^{\beta})(\alpha e^{\lambda x^{\beta}} + 1 - \alpha) + \lambda\beta(1-\alpha)x^{\beta} = 0$. This simplifies to $\alpha(\beta - 1 - \lambda\beta x^{\beta})e^{\lambda x^{\beta}} + (1-\alpha)(\beta-1) = 0$. Hence, the critical values of $f(x)$ are $x = 0$ or $x = x_0$ where $k(x_0) = 0$. Consider the derivative of $k(x)$ as $k'(x) = \alpha\beta e^{x^{\beta}}x^{\beta-1}(1 + x^{\beta}\beta)$. Clearly $k'(x) > 0$ for all $x > 0$. Therefore, $k(x)$ is strictly increasing. Now assume $\beta \leq 1$. Since $k(0) = 1 - \beta \geq 0$, this implies that $x = 0$ is the only critical values of $f(x)$. Also, $\lim_{x \to 0} f(x) = \infty$ if $\beta < 1$ and $(\alpha - 1)\log\alpha$ if $\beta = 1$. Hence, the mode is at $x = 0$. Now assume $\beta > 1$. The fact that $k(0) = 1 - \beta < 0$ and $k(x)$ is strictly increasing, implies that $k(x)$ has a unique solution at $x = x_0$. Furthermore, when $\beta > 1$, $\lim_{x \to 0} f(x) = 0$ and therefore, $x = 0$ is not a modal point. This completes the proof.                                                                              □

## 4.1 Moment and Moment Generating Function

In this subsection, we will derive the $r$-th moments and the moment generating function for the APWQ distribution.

If $0 \leq \alpha \leq 2, \alpha \neq 1$ and From (10), it is easy to obtain the $r$-th moment of APWQ as

$$E(X^r) = \frac{\Gamma(1 + r/\beta)}{\lambda^{r/\beta}\log(\alpha)}\sum_{k=0}^{\infty}\sum_{j=0}^{k}\binom{k}{j}\frac{(-1)^{k+j}(\alpha - 1)^{k+1}}{(j+1)^{1+r/\beta}}. \tag{12}$$

Similarly, if $\alpha > 2$, the $r$-th moment of APWQ can be obtained from (11) as

$$E(X^r) = \frac{\Gamma(1 + r/\beta)}{\lambda^{r/\beta}\log(\alpha)}\sum_{k=0}^{\infty}\frac{(1 - 1/\alpha)^{k+1}}{(k+1)^{1+r/\beta}} \tag{13}$$

Also, the moment generating function for $0 \leq \alpha \leq 2, \alpha \neq 1$, can be written as

$$M_x(t) = \frac{1}{\log(\alpha)}\sum_{k=0}^{\infty}\sum_{j=0}^{k}\sum_{m=0}^{\infty}\binom{k}{j}\frac{(-1)^{k+j}(\alpha - 1)^{k+1}t^m}{m!(j+1)^{1+m/\beta}\lambda^{m/\beta}}\Gamma(1 + m/\beta).$$

Similarly, the moment generating function for $\alpha > 2$, takes the form

$$M_x(t) = \frac{1}{\log(\alpha)}\sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\frac{(1 - 1/\alpha)^{k+1}t^m\Gamma(1 + m/\beta)}{m!(k+1)^{1+m/\beta}\lambda^{m/\beta}}$$

Remark 3, Theorem 2 and Eqs. (12) and (13) are used to obtain the mean, median, mode, variance, skewness and kurtosis for the APWQ distribution. The median is

**Table 2** Mean, median, mode, variance, skewness, and kurtosis of APWQ for $\lambda = 1$ and various values of $\alpha$ and $\beta$

| $\alpha$ | $\beta$ | Mean | Median | Mode | Variance | Skewness | Kurtosis |
|---|---|---|---|---|---|---|---|
| 0.5 | 0.5 | 1.499 | 0.777 | 0 | 4.275 | 2.006 | 6.510 |
| | 0.8 | 1.352 | 0.854 | 0 | 2.223 | 1.985 | 7.808 |
| | 1.5 | 1.023 | 0.919 | 0.655 | 0.419 | 0.899 | 3.902 |
| | 3 | 0.957 | 0.959 | 0.976 | 0.107 | 0.04 | 2.699 |
| | 5 | 0.959 | 0.975 | 1.014 | 0.043 | $-0.372$ | 2.997 |
| 0.8 | 0.5 | 1.310 | 0.563 | 0 | 3.770 | 2.215 | 7.635 |
| | 0.8 | 1.184 | 0.699 | 0 | 1.936 | 2.203 | 9.115 |
| | 1.5 | 0.941 | 0.826 | 0.534 | 0.389 | 1.015 | 4.223 |
| | 3 | 0.914 | 0.909 | 0.907 | 0.106 | 0.126 | 2.717 |
| | 5 | 0.931 | 0.944 | 0.975 | 0.044 | $-0.293$ | 2.916 |
| 1.5 | 0.5 | 1.080 | 0.356 | 0 | 3.155 | 2.538 | 9.551 |
| | 0.8 | 0.989 | 0.525 | 0 | 1.616 | 2.510 | 11.156 |
| | 1.5 | 0.837 | 0.709 | 0.393 | 0.353 | 1.177 | 4.717 |
| | 3 | 0.856 | 0.842 | 0.813 | 0.104 | 0.245 | 2.760 |
| | 5 | 0.894 | 0.902 | 0.921 | 0.045 | $-0.184$ | 2.820 |
| 3 | 0.5 | 0.868 | 0.208 | 0 | 2.585 | 2.932 | 12.21 |
| | 0.8 | 0.812 | 0.374 | 0 | 1.336 | 2.861 | 13.79 |
| | 1.5 | 0.733 | 0.592 | 0.270 | 0.320 | 1.359 | 5.340 |
| | 3 | 0.794 | 0.770 | 0.708 | 0.103 | 0.376 | 2.836 |
| | 5 | 0.852 | 0.855 | 0.858 | 0.047 | $-0.064$ | 2.735 |
| 5 | 0.5 | 0.741 | 0.137 | 0 | 2.240 | 3.236 | 14.502 |
| | 0.8 | 0.705 | 0.288 | 0 | 1.170 | 3.124 | 15.972 |
| | 1.5 | 0.664 | 0.515 | 0.200 | 0.298 | 1.494 | 5.847 |
| | 3 | 0.749 | 0.718 | 0.632 | 0.103 | 0.470 | 2.909 |
| | 5 | 0.821 | 0.820 | 0.810 | 0.048 | 0.023 | 2.688 |

obtained by setting $q = 0.5$ in Remark 3. The mean $\mu$ is obtained by setting $r = 1$ in Eqs. 12 or 13 based on the value of $\alpha$. The variance $\sigma^2$, skewness $\gamma_1$ and kurtosis $\gamma_2$ are obtained using the formulas $\sigma^2 = E(X^2) - \mu^2$, $\gamma_1 = E[(X - \mu)/\sigma]^3$ and $\gamma_2 = E[(X - \mu)/\sigma]^4$. These values are reported in Table 2 for various values of $\alpha$ and $\beta$ where the scale parameter $\lambda = 1$. From Table 2, it is noted that for fixed $\beta$ and $\lambda$, the mean, median and mode of APWQ are decreasing function of $\alpha$, and the skewness is increasing function of $\alpha$. Also, for fixed $\alpha$ and $\lambda$, the median is an increasing function of $\beta$, the mode is an increasing function of $\beta > 1$, while the variance and skewness are decreasing function of $\beta$. Also, Table 2 shows that the APWQ is a flexible distribution. It can be left skewed, right skewed or approximately symmetric. Furthermore, it can be platykurtic (kurtosis $< 3$) or leptokurtic (kurtosis $> 3$).

## 4.2 Shannon Entropy

Using (7) and (9), the Shannon entropy, $\eta_X$, for $0 \leq \alpha \leq 2, \alpha \neq 1$ is given by

$$\eta_X = \log\left(\frac{\sqrt{\alpha}\log(\alpha)}{(\alpha-1)\lambda\beta}\right) - \int_0^\infty \{I_1 - I_2\} f(x)dx \tag{14}$$

where $I_1 = (\beta - 1)\log(x)$ and $I_2 = \lambda x^\beta$. Now

$$\int_0^\infty I_1 f(x)dx = \frac{\lambda\beta(\beta-1)}{\log(\alpha)} \sum_{k=0}^\infty \sum_{j=0}^k \binom{k}{j}(-1)^{k+j}(\alpha-1)^{k+1}$$
$$\times \int_0^\infty x^{\beta-1}e^{-\lambda(j+1)x^\beta}\log(x)dx. \tag{15}$$

On using the $\int_0^\infty e^{-ax}\log x\,dx = -\frac{1}{a}(C + \log a)$, where $C$ is the Euler constant, (15) can be written as

$$\int_0^\infty I_1 f(x)dx = \frac{\beta-1}{\log(\alpha)} \sum_{k=0}^\infty \sum_{j=0}^k \binom{k}{j}\frac{(-1)^{k+j-1}}{\beta(j+1)}(\alpha-1)^{k+1}\left[C + \log(\lambda(j+1))\right]. \tag{16}$$

Similarly,

$$\int_0^\infty I_2 f(x)dx = \frac{1}{\log(\alpha)} \sum_{k=0}^\infty \sum_{j=0}^k \binom{k}{j}(-1)^{k+j}\frac{(\alpha-1)^{k+1}}{(j+1)^2}. \tag{17}$$

From (14), (16) and (17), $\eta_X$ reduces to

$$\eta_X = \log\left(\frac{\sqrt{\alpha}\log(\alpha)}{(\alpha-1)\lambda\beta}\right) - \frac{\beta-1}{\log(\alpha)} \sum_{k=0}^\infty \sum_{j=0}^k \binom{k}{j}\frac{(-1)^{k+j+1}(\alpha-1)^{k+1}}{j+1}$$
$$\times \left[\frac{C + \log(\lambda(j+1))}{\beta} + \frac{1}{(\beta-1)(j+1)}\right].$$

Using similar approach, the Shannon entropy, for $\alpha > 2$ is given by

$$\eta_X = \log\left(\frac{\sqrt{\alpha}\log(\alpha)}{(\alpha-1)\lambda\beta}\right) + \frac{\beta-1}{\log(\alpha)} \sum_{k=0}^\infty \frac{(1-1/\alpha)^{k+1}}{k+1}\left[\frac{C+\log(\lambda(k+1))}{\beta} + \frac{1}{(\beta-1)(k+1)}\right].$$

## 4.3 Mean Residual Life and Mean Waiting Time

Let $X$ be a continuous random variable. The mean residual life is the expected additional lifetime that a component has survived after a fixed time point $t$. The mean residual life function, say $\mu(t)$, is given by

$$\mu(t) = E(T - t | T > t) = \frac{1}{S(t)} \int_t^\infty x f(x) dx - t,$$

where

$$\int_t^\infty x f(x) dx = \frac{\lambda^{-1/\beta}}{\log(\alpha)} \sum_{k=0}^\infty \sum_{j=0}^k \binom{k}{j} (-1)^{k+j} (\alpha - 1)^{k+1} \frac{\Gamma(\lambda(j+1)t^\beta, 1 + 1/\beta)}{(j+1)^{1+1/\beta}},$$

where $\Gamma(a, b)$ is the upper incomplete gamma function and $0 \le \alpha \le 2, \alpha \ne 1$. When $\alpha > 2$ we have

$$\int_t^\infty x f(x) dx = \frac{\lambda^{-1/\beta}}{\log(\alpha)} \sum_{k=0}^\infty (1 - 1/\alpha)^{k+1} \frac{\Gamma(\lambda(k+1)t^\beta, 1 + 1/\beta)}{(k+1)^{1+1/\beta}}$$

The mean waiting time represents the waiting time elapsed since the failure of an object on condition that this failure had occurred in the interval $[0, t]$. The mean waiting time of $X$, say $\bar{\mu}(t)$, is defined by

$$\bar{\mu}(t) = t - \frac{m(t)}{F(t)} \tag{18}$$

where $F(t)$ is the CDF given by (8) and $m(t)$ is the first incomplete moment given by

$$m(t) = \int_0^t x f(x) dx$$
$$= \frac{\lambda^{-1/\beta}}{\log(\alpha)} \sum_{k=0}^\infty \sum_{j=0}^k \binom{k}{j} (-1)^{k+j} (\alpha - 1)^{k+1} \frac{\gamma\{\lambda(j+1)t^\beta, 1 + 1/\beta\}}{(j+1)^{1+1/\beta}} \tag{19}$$

where $\gamma(a, b)$ is the lower incomplete gamma function and $0 \le \alpha \le 2, \alpha \ne 1$. Substituting (8) and (19) in (18), $\bar{\mu}(t)$ can be written as

$$\bar{\mu}(t) = t - \lambda^{-1/\beta} \sum_{k=0}^\infty \sum_{j=0}^k \binom{k}{j} (-1)^{k+j} (\alpha - 1)^{k+1} \frac{\gamma\{\lambda(j+1)t^\beta, 1 + 1/\beta\}}{(j+1)^{1+1/\beta} \log(\alpha + (1 - \alpha)e^{-\lambda t^\beta})}.$$

Similarly, $\bar{\mu}(t)$ in the case of $\alpha > 2$ can be written as

$$\bar{\mu}(t) = t - \lambda^{-1/\beta} \sum_{k=0}^\infty (1 - 1/\alpha)^{k+1} \frac{\gamma\{\lambda(k+1)t^\beta, 1 + 1/\beta\}}{(k+1)^{1+1/\beta} \log(\alpha + (1 - \alpha)e^{-\lambda t^\beta})}.$$

## 5 Estimation and Applications

Let $x_1, x_2, \ldots, x_n$ be a random sample from APWQ. The log-likelihood function is given by

$$\ell(\alpha, \lambda, \beta) = n \log(\alpha - 1) - n \log(\log(\alpha)) + n \log(\lambda\beta) + (\beta - 1) \sum_{i=1}^{n} \log(x_i)$$

$$-\lambda \sum_{i=1}^{n} x_i^{\beta} - \sum_{i=1}^{n} \log[1 + (\alpha - 1)\varphi_i], \qquad (20)$$

where $\varphi_i = 1 - e^{-\lambda x_i^{\beta}}$.

Therefore, the MLE's of $\alpha, \lambda$ and $\beta$ can be computed by maximizing the log-likelihood function in (20). We used the routine OPTIM which is available in the R software. Next, the APWQ distribution is used to model different types of cancer data sets including complete and censored data.

### 5.1 Complete Data

The data set represents the survival times of 121 patients with breast cancer obtained from a large hospital in a period from 1929 to 1938 [14]. This data set has recently been studied by [21]. The data are:

0.3, 0.3, 4.0, 5.0, 5.6, 6.2, 6.3, 6.6, 6.8, 7.4, 7.5, 8.4, 8.4, 10.3,11.0, 11.8, 12.2, 12.3, 13.5, 14.4, 14.4, 14.8, 15.5, 15.7, 16.2, 16.3, 16.5, 16.8, 17.2, 17.3, 17.5,17.9, 19.8, 20.4, 20.9, 21.0, 21.0, 21.1, 23.0, 23.4, 23.6, 24.0, 24.0, 27.9, 28.2, 29.1, 30.0, 31.0,31.0, 32.0, 35.0, 35.0, 37.0, 37.0, 37.0, 38.0, 38.0, 38.0, 39.0, 39.0, 40.0, 40.0, 40.0, 41.0, 41.0,41.0, 42.0, 43.0, 43.0, 43.0, 44.0, 45.0, 45.0, 46.0, 46.0, 47.0, 48.0, 49.0, 51.0, 51.0, 51.0, 52.0,54.0, 55.0, 56.0, 57.0, 58.0, 59.0, 60.0, 60.0, 60.0, 61.0, 62.0, 65.0, 65.0, 67.0, 67.0, 68.0, 69.0,78.0, 80.0,83.0, 88.0, 89.0, 90.0, 93.0, 96.0, 103.0, 105.0, 109.0, 109.0, 111.0, 115.0, 117.0, 125.0,126.0, 127.0, 129.0, 129.0, 139.0, 154.0.

The APWQ distribution is fitted to the data set and compared with several other competitive models namely: McDonald Weibull (Mc-W) [6], Beta Weibull (BW) [13], Modified Weibull (MW) [4], Marshall-Olkin Weibull (MOW) [17] and Zografos-Balakrishnan log-logistic (ZBLL) [25].

Table 3 lists the MLEs (and the corresponding standard errors in parentheses) of the parameters, negative likelihood values $[-\ell(\hat{\theta})]$, Kolmogorov–Smirnov (K–S) test and the $p$ value for the K–S statistics for all fitted models. From Table 3, it is observed that the APWQ distribution has the lowest values of $[-\ell(\hat{\theta})]$ and K–S and the largest $p$ value for the K–S statistics, which implies that the APWQ distribution provides the best fit among all fitted distributions followed by MOW distribution. Figures 3a displays the histogram and the fitted APWQ density for the data set. Also, the plots of the fitted APWQ survival and the empirical survival functions for the data set are

**Table 3** The MLEs (standard errors in parentheses), and the goodness of fit statistics for the breast cancer data

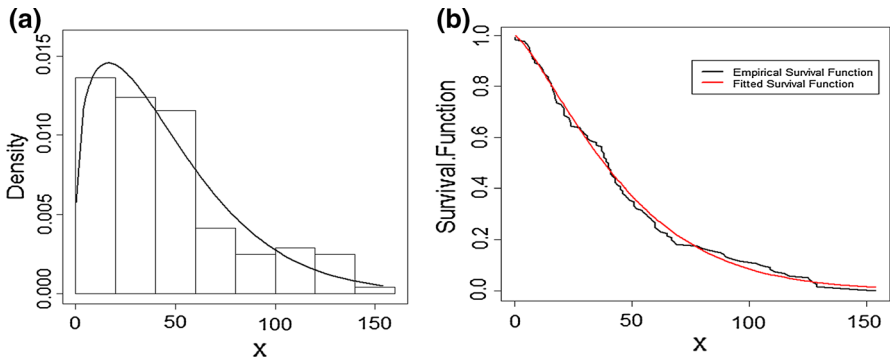| Model | Estimates | | | | | Statistics | | |
|---|---|---|---|---|---|---|---|---|
| | $\hat{\alpha}$ | $\hat{\beta}$ | $\hat{\lambda}$ | $\hat{b}$ | $\hat{k}$ | $-\ell(\hat{\theta})$ | $K-S$ | $p$ value |
| Mc-W | 30.72483 | 0.05285 | 2.41657 | 40.54543 | 14.9355 | 594.700 | 0.11718 | 0.0721 |
| | (18.7724) | (0.014) | (0.7272) | (28.263) | (12.2245) | | | |
| BW | 34.63665 | 0.18781 | 0.17746 | 11.67641 | – | 590.968 | 0.11079 | 0.10256 |
| | (17.6874) | (0.0434) | (0.2146) | (7.349) | | | | |
| APWQ | 0.87231 | 1.286741 | 0.006755 | – | – | 579.020 | 0.05025 | 0.92004 |
| | (1.505) | (0.2531) | (0.0099) | | | | | |
| MW | 0.02158 | 0.99997 | 0.00001 | – | – | 585.128 | 0.11213 | 0.09541 |
| | (0.00139) | (2.593) | (0.00139) | | | | | |
| MOW | 33.95266 | 0.54287 | 0.26894 | – | – | 580.195 | 0.0536 | 0.87777 |
| | (16.429) | (0.422) | (0.964) | | | | | |
| ZBLL | 1.5738 | 1.6204 | 19.077 | – | – | 608.990 | 0.1014 | 0.16603 |
| | (0.09626) | (0.118) | (0.2103) | | | | | |

**(a)**



**(b)**



**Fig. 3** **a** Histogram and the fitted APWQ distribution. **b** Fitted APWQ survival and empirical survival functions for the first data set

displayed in Fig. 3b. It is clear that these plots show that APWQ provides good fit to the data set and this supports the results in Table 3.

### 5.2 Censored Data

Censored data are very common in lifetime applications. Some mechanisms of censoring are identified in literature such as type I and II censoring. The fact that APWQ has closed form survival function advantages the distribution to be used in analyzing lifetime data in the presence of censoring. Consider a data set $D = (x; r)$, where $x = (x_1, x_2, x_3, \ldots, x_n)$ are the observed failure times and $r_i = (r_1, \ldots, r_n)$ are the censored failure times where $r_i$ is equal to 1 if a failure is observed and 0 otherwise. Suppose that the data are independently and identically distributed follows a distribution with probability density and survival functions $f(x, \theta)$ and $S(x, \theta)$ respectively. Then the likelihood function for parameters $\theta = (\alpha, \lambda, \beta)^T$ can be written as

$$L(D; \theta) = \prod_{i=1}^{n} [f(x_i, \theta)]^{r_i} [S(x_i, \theta)]^{1-r_i}.$$

For the APWQ distribution, the log-likelihood function is given by

$$\ell = r \log\left(\frac{(\alpha-1)\lambda\beta}{\log(\alpha)}\right) + \sum_{i=1}^{n} r_i \left((\beta-1)\log x_i - \lambda x_i^{\beta} - \log(1 + (\alpha-1)\varphi_i)\right)$$

$$+ \sum_{i=1}^{n} (1 - r_i) \left\{\log\left[\log(\alpha) - \log(1 + (\alpha-1)\varphi_i)\right] - \log(\log(\alpha))\right\} \qquad (21)$$

where $r = \sum_{i=1}^{n} r_i$ and $\varphi_i$ is defined in (20). The log likelihood function in (21) can be maximized numerically in order to obtain the ML estimates. The routine OPTIM which is available in the R software can be used.

434

Ann. Data. Sci. (2018) 5(3):421–436

**Table 4** The MLEs (standard errors in parentheses), and the corresponding AIC and BIC values for second censored data set

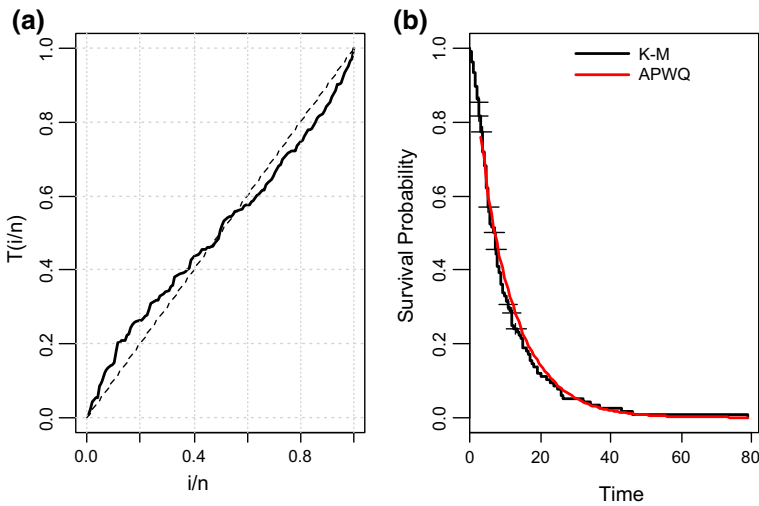| Model | Estimates | | | AIC | BIC |
|---|---|---|---|---|---|
| APWQ($\alpha$, $\beta$, $\lambda$) | 6.0377 | 1.2969 | 0.0269 | 837.7477 | 843.5723 |
| | (1.6640) | (0.20793) | (0.0250) | | |
| MOW($\alpha$, $\beta$, $\lambda$) | 0.0682 | 1.6351 | 0.0282 | 843.4338 | 852.1937 |
| | (0.1623) | (0.3509) | (0.0278) | | |
| W($\beta$, $\lambda$) | – | 1.0535 | 10.2008 | 848.8294 | 854.6693 |
| | | (0.06813) | (0.8921) | | |



**Fig. 4** **a** TTT plot **b** fitted survival curve of APWQ distribution

We consider a censored data set that contains remission times for bladder cancer patients. The data set has 137 observations with 9 censored. More details about the data can be found in Lee and Wang [15]. The TTT plot in Fig. 4a is concave then convex which gives an indication of upside down bathtub failure rate. The distribution fits are given in Table 4. From the table, we can see that APWQ distribution has the lowest Akaike information criterion (AIC) and Bayesian information criterion (BIC) values as compared to other fitted models.

The survival curve of the fitted APWQ distribution given in Fig. 4b fits the Kaplan Meier curve well.

## 6 Conclusions

In this paper, a method for generating family of distributions is proposed based on the APT family proposed recently by Mahdavi and Kundu [16]. The proposed method can produce a flexible hazard rate functions. Some general properties of the proposed

family are studied. A member of the proposed family, APWQ distribution is studied in details. The APWQ distribution is the generalization of the Weibull distribution with attractive shape flexibilities for both the density and the hazard rate functions. In fact, the density function can be left-skewed, right-skewed or about symmetric. The hazard rate function possesses an IFR, DFR, BT or UBT shapes. Real data sets are used to show the applicability of the APWQ distribution to complete as well as censored data sets. The fact that APWQ has only three parameters with closed form CDF and at the same time possesses several types of hazard rate shapes; make this distribution an attractive choice to be used in various filed of studies including cancer research.

# References

1. Alzaatreh A, Famoye F, Lee C (2014a) The gamma-normal distribution: properties and applications. Comput Stat Data Anal 69:67–80
2. Alzaatreh A, Famoye F, Lee C (2014b) T-normal family of distribution: a new approach to generalize the normal distribution. J Stat Distrib Appl 1:1–16
3. Alzaatreh A, Lee C, Famoye F (2013) A new method for generating families of continuous distributions. Metron 71:63–79
4. Ammar M, Mazen Z (2009) Modified Weibull distribution. Appl Sci 11:123–136
5. Cordeiro GM, Edwin MM, Lemonte AJ (2014) The exponential-Weibull lifetime distribution. J Stat Comput Simul 84:2592–2606
6. Corderio GM, Hashimoto EH, Ortega EMM (2012) The McDonald Weibull model. Statistics 48:256–278
7. Eugene N, Lee C, Famoye F (2002) The beta-normal distribution and its applications. Commun Stat Theory Methods 31:497–512
8. Jalmar MF, Edwin MM, Cordeiro GM (2008) A generalized modified Weibull distribution for lifetime modeling. Comput Stat Data Anal 53:450–462
9. Jones MC (2009) Kumaraswamy's distribution: a beta type distribution with tractability advantages. Stat Methodol 6:70–81
10. Johnson NL (1949) Systems of frequency curves generated by methods of translation. Biometrika 36:149–176
11. Johnson NL, Kotz S, Balakrishnan N (1994) Continuous univariate distributions, vol 1, 2nd edn. John Wiley and Sons Inc., New York
12. Lee C, Famoye F, Alzaatreh A (2013) Methods for generating families of continuous distribution in the recent decades. Wiley Interdiscip Rev Comput Stat 5:219–238
13. Lee C, Famoye F, Olumolade O (2007) Beta Weibull distribution: Some properties and applications to censored data. J Modern Appl Stat Methods 6:173–186
14. Lee ET (1992) Statistical methods for survival data analysis. John Wiley, New York
15. Lee ET, Wang JW (2003) Statistical methods for survival data analysis, 3rd edn. John Wiley, New York
16. Mahdavi A, Kundu D (2017) A new method for generating distributions with an application to exponential distribution. Commun Stat Theory Methods 46(13):6543–6557
17. Marshall AN, Olkin I (1997) A new method for adding a parameter to a family of distributions with applications to the exponential and Weibull families. Biometrica 84:641–652
18. Mudholkar GS, Srivastava DK, Friemer M (1995) The exponentiated Weibull family: a reanalysis of the bus-motor-failure data. Technometrics 37:436–445
19. Mudholkar GS, Srivastava DK, Kollia GD (1996) A generalization of the Weibull distribution with application to the analysis of survival data. J Am Stat Assoc 91:1575–1583
20. Pearson K (1895) Contributions to the mathematical theory of evolution. II. Skew variation in homogeneous material. Philos Trans R Soc Lond A 186:343–414
21. Ramos MWA, Cordeiro GM, Marinho PRD, Dias CRB, Hamedani GG (2013) The Zografos-Balakrishnan log-logistic distribution: Properties and applications. J Stat Theory Appl 12:225–244

22. Tahir M, Zubair M, Cordeiro G, Alzaatreh A, Mansoor M (2016) The Poisson-X family of distributions. J Stat Comput Simul 86(14):2901–2921. https://doi.org/10.1080/00949655.2016.1138224
23. Tukey JW (1960) The practical relationship between the common transformations of percentages of counts and amounts. Technical Report 36. Statistical Techniques Research Group, Princeton University, Princeton, NJ
24. Xie M, Lai CD (1995) Reliability analysis using an additive Weibull model with bathtub-shaped failure rate function. Reliab Eng Syst Saf 52:87–93
25. Zografos K, Balakrishnan N (2009) On families of beta- and generalized gamma-generated distributions and associated inference. Stat Methodol 6:344–362